



LEGAL NOTICE:

© **Copyright 2008 to 2022 NVM Express®, Inc. ALL RIGHTS RESERVED.**

This Technical Proposal is proprietary to the NVM Express, Inc. (also referred to as "Company") and/or its successors and assigns.

NOTICE TO USERS WHO ARE NVM EXPRESS, INC. MEMBERS: Members of NVM Express, Inc. have the right to use and implement this Technical Proposal subject, however, to the Member's continued compliance with the Company's Intellectual Property Policy and Bylaws and the Member's Participation Agreement.

NOTICE TO NON-MEMBERS OF NVM EXPRESS, INC.: If you are not a Member of NVM Express, Inc. and you have obtained a copy of this document, you only have a right to review this document or make reference to or cite this document. Any such references or citations to this document must acknowledge NVM Express, Inc. copyright ownership of this document. The proper copyright citation or reference is as follows: "© 2008 to 2022 NVM Express, Inc. ALL RIGHTS RESERVED." When making any such citations or references to this document you are not permitted to revise, alter, modify, make any derivatives of, or otherwise amend the referenced portion of this document in any way without the prior express written permission of NVM Express, Inc. Nothing contained in this document shall be deemed as granting you any kind of license to implement or use this document or the specification described therein, or any of its contents, either expressly or impliedly, or to any intellectual property owned or controlled by NVM Express, Inc., including, without limitation, any trademarks of NVM Express, Inc.

LEGAL DISCLAIMER:

THIS DOCUMENT AND THE INFORMATION CONTAINED HEREIN IS PROVIDED ON AN "AS IS" BASIS. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, NVM EXPRESS, INC. (ALONG WITH THE CONTRIBUTORS TO THIS DOCUMENT) HEREBY DISCLAIM ALL REPRESENTATIONS, WARRANTIES AND/OR COVENANTS, EITHER EXPRESS OR IMPLIED, STATUTORY OR AT COMMON LAW, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, VALIDITY, AND/OR NONINFRINGEMENT.

All product names, trademarks, registered trademarks, and/or servicemarks may be claimed as the property of their respective owners.

The NVM Express® design mark is a registered trademark of NVM Express, Inc.

NVM Express
c/o VTM, Inc.
3855 SW 153rd Drive
Beaverton, OR 97003
USA
info@nvmexpress.org

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

NVM Express® Technical Proposal for New Feature

Technical Proposal ID	TP 6027 - Reset Behavior Clarifications
Change Date	2022.03.02
Builds on Specification	NVM Express Management Interface 1.2b

Technical Proposal Author(s)

Name	Company
Yoni Shternhell	WDC
Myron Loewen	Intel
Austin Bolen	Dell EMC

This TP modifies the MI specification with implicit reset naming and removes ambiguous behavior for different form factor signals

Revision History

Revision Date	Change Description
10/26/2020	Initial Draft
11/02/2020	Added SFF 1009 reference
11/09/2020	Remove modification to the VPD section
12/09/2020	Modifications to the SMBRST section
01/04/2021	Modifications to the SMBRST section and 2021 editorial changes
02/10/2021	Modify the SMBus Reset section with: <ul style="list-style-type: none">- NVMe-MI should not specify reset behavior for ARP and EID as they are SMBus/MCTP properties, respectively- SMBus frequency and MTU default values on SMBus Clock Low Reset, SMBus Power On Reset, Controller Level Reset, etc- SMBus clock low timings change from 10ms to 50ms
02/22/2021	Added definition to SMBus Reset
04/19/2021	Updated for NVMe-MI 1.2, added summary for change document
04/23/2021	Fixed the navigation pane. Removed redundant Section 4.2 changes.
05/03/2021	Added a note in the Operations Supported During NVM Subsystem Power States
05/10/2021	Added clarifications for more ambiguous uses of "reset".
05/17/2021	Updates from workgroup call on 5/17/2021.

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Revision Date	Change Description
05/24/2021	<ul style="list-style-type: none"> Added that Controller Level Reset sets status bits back to the default in the in-band tunneling mechanism and Management Endpoint Reset sets status bits back to the default in the out-of-band mechanism. Updates from workgroup call on 5/24/2021.
06/06/2021	<ul style="list-style-type: none"> Resolved many comments. Updates from workgroup call on 6/7/2021.
06/07/2021	<ul style="list-style-type: none"> Resolved more comments. Updates from workgroup call on 6/14/2021.
06/21/2021	<ul style="list-style-type: none"> Resolved more comments. Updates from workgroup call on 6/21/2021.
07/12/2021	<ul style="list-style-type: none"> Resolved more comments. Updates from workgroup call on 7/12/2021.
07/19/2021	<ul style="list-style-type: none"> Cleanup for phase 2 exit.
07/20/2021	<ul style="list-style-type: none"> Updates for phase 3.
08/30/2021	<ul style="list-style-type: none"> Update description for changes document. Clarified how status bits are set to their reset value due to Controller Level Reset or Management Endpoint Reset. Reverted to 1.2 text for impacts due firmware update.
10/14/2021	<ul style="list-style-type: none"> Updates based on member review feedback.
10/23/2021	<ul style="list-style-type: none"> More updates based on member review feedback.
10/29/2021	<ul style="list-style-type: none"> More updates based on member review feedback.
11/07/2021	<ul style="list-style-type: none"> More updates based on member review feedback.
11/08/2021	<ul style="list-style-type: none"> More updates based on member review feedback.
02/02/2022	<ul style="list-style-type: none"> Integrated
02/28/2022	<ul style="list-style-type: none"> Updated section 8.1 Figure 146 Note 4 to reference section 8.3.2

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Description for NVMe-MI Changes Document

Feature Enhancement:

- Clarify and improve the effects of resets in NVMe-MI.
- Added a reference to SNIA SFF-TA-1009 Enterprise and Datacenter Standard Form Factor Pin and Signal Specification.
- Clarified in the Operations Supported During NVM Subsystem Power States table that SMBus Reset or PCIe Reset may prevent MCTP access.
- **Incompatible change**
 - Some SMBus resets are no longer required to maintain ARP assigned address.
 - Changed the type of reset required to return to normal operation when the Reset Not Required bit is cleared to '0' from Controller Level Reset to NVM Subsystem Reset.
 - The reset value of the Controller Enable Change Occurred bit was changed from 0 to HwInit for the in-band tunneling mechanism.
 - Status bits in the in-band tunneling mechanism now have an instance per Controller instead of a single instance shared among all Controllers in the NVM Subsystem.
- **New requirement**
 - SMBus resets are now required to reset SMBus/I2C MTU and frequency.
 - SMBus/I2C Management Endpoints are now required to support SMBus clock-low recovery.
 - Added PCIe Reset requirements.
 - SMBus Reset timing is now the same as SMBus power on timing.
 - SMBus Reset shall cause the SMBus/I2C port to attempt to generate a STOP condition within 5 ms from the assertion of SMBus Reset.
 - Specified what is reset by an NVM Subsystem Reset.
 - Specified that a Management Endpoint Reset aborts outstanding Command Messages and discards outstanding Control Primitives.
- References:
 - Technical Proposal 6027.

Description of Specification Changes

Markup Conventions:

Black:	Unchanged (however, hot links are removed)
Red Strikethrough:	Deleted
<u>Blue Underline:</u>	New
Blue Highlighted:	TBD values, anchors, and links to be inserted in new text.
<Green Bracketed>:	Notes to editor

1.7 Conventions

Modify a portion of Section 1.7 (Conventions) as follows:

...

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Hardware Initialized (HwInit) – The ~~default~~ value is dependent on Responder and system configuration. For the out-of-band mechanism, ~~the value is initialized at by an NVM Subsystem Reset (e.g., by an expansion ROM, or in the case of integrated devices, by a platform BIOS).~~ For the in-band tunneling mechanism, the value is initialized by a Controller Level Reset (refer to the NVM Express Base Specification).

Reset – For the out-of-band mechanism, this column indicates the value the field is initialized to by an NVM Subsystem Reset. For the in-band tunneling mechanism, this column indicates the value the field is initialized to by a Controller Level Reset (refer to the NVM Express Base Specification).

Modify a portion of Section 1.8 (Definitions) as follows:

...

1.8. ~~TBD~~ Management Endpoint Reset

A mechanism used to reset a Management Endpoint in an NVMe Storage Device or NVMe Enclosure. For more information, see section ~~8.3.3~~.

1.8. ~~TBD~~ PCIe Reset

A mechanism used to reset a PCIe VDM Management Endpoint in an NVMe Storage Device or NVMe Enclosure. For more information, see section ~~8.3.TBD~~.

1.8. ~~TBD~~ SMBus Reset

A mechanism used to reset the SMBus/I2C elements in an NVMe Storage Device or NVMe Enclosure. For more information, see section ~~8.3.4~~.

Add the following reference entry to Section 1.1.1 References:

1.11 References

...

SNIA SFF-TA-1008 Enterprise and Datacenter ~~3" SSD~~ Device Form Factor (E3) Specification. Available from <https://www.snia.org>.

SNIA SFF-TA-1009 Enterprise and Datacenter Standard Form Factor Pin and Signal Specification. Available from <https://www.snia.org>.

...

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Modify a portion of Section 4.2 (Out-of-Band Message Servicing Model) as follows:

4.2 Out-of-Band Message Servicing Model

...

Idle: A Command Slot is idle when it is not in the Receive, Process, or Transmit state. This is the default initial state of the command servicing state machine (e.g., following a Management Endpoint Reset (refer to section 8.3.3)). Command servicing transitions from the Idle state to the Receive state when the first MCTP packet of a Command Message is received (i.e., an MCTP packet with the SOM bit in the MCTP packet header set to '1' and the Message Type set to 4h).

Modify Figure 42 (Get State Control Primitive Success Response Fields) as follows:

4.2.1.4 Get State

...

Figure 42: Get State Control Primitive Success Response Fields

Bytes	Description		
07:06	Control Primitive Specific Response (CPSR): This field is used to return Management Endpoint Control Primitive-specific status. A Management Endpoint Reset of the corresponding Management Endpoint shall clear this field to 0h.		
	Bits	Command Slot Specific ¹	Description
	15	Yes	Pause Flag (PFLG): This bit indicates whether or not the Command Slot is paused. This bit set to '1' indicates the Command Slot is paused. This bit cleared to '0' indicates the Command Slot is not paused. While the Pause Flag is set, the Management Endpoint disables the timeout waiting for packet timer (refer to section 4.2.1.1) for the Command Slot and does not transmit responses to Command Messages.
	14	No	NVM Subsystem Reset Occurred (NSSRO): This bit indicates when an NVM Subsystem Reset occurs while main power is applied. This bit is set to '1' if the last occurrence of an NVM Subsystem Reset occurred while main power was applied to the NVM Subsystem. This bit is cleared to '0' following a power cycle and following a Get State Control Primitive with the CESF bit set to '1'.
	13	No	Bad Packet or Other Physical Layer (BPOPL): This bit is set to '1' if a packet sent to the Management Endpoint failed a transport specific packet integrity check since the last time Get State Control Primitive was processed with the CESF bit set to '1'.
	12	No	Bad, Unexpected, or Expired Message Tag (BUENT): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.
	11	No	Out-of-Sequence Packet Sequence Number (OSPSN): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.
	10	No	Unexpected Middle or End of Packet (UMEP): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Figure 42: Get State Control Primitive Success Response Fields

Bytes	Description											
	09	No	Incorrect Transmission Unit (ITU): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	08	No	Unknown Destination ID (UDSTID): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	07	No	Bad Header Version (BHVS): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	06	No	Unsupported Transmission Unit (UTUNT): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	05	No	Timeout Waiting for a Packet (WPTT): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	04	No	Bad Message Integrity Check Error (BMICE): This bit is set to '1' if the Management Endpoint detected an error of this type (refer to the MCTP Base Specification) since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	03	No	Command Message to non-Idle Command Slot (CMNICS): This bit is set to '1' if the Management Endpoint discarded one or more Command Messages due to overlapping Command Messages to a Command Slot since the last time Get State Control Primitive was processed with the CESF bit set to '1'.									
	02		Reserved									
	01:00	Yes	Slot Command Servicing State (SSTA): This field indicates the current command servicing state of the Command Slot. An implementation may choose to indicate only the Idle and Process states in this field. Refer to Figure 33. <table><tr><th>Value</th><th>Description</th></tr><tr><td>0h</td><td>Idle</td></tr><tr><td>1h</td><td>Receive</td></tr><tr><td>2h</td><td>Process</td></tr><tr><td>3h</td><td>Transmit</td></tr></table>	Value	Description	0h	Idle	1h	Receive	2h	Process	3h
Value	Description											
0h	Idle											
1h	Receive											
2h	Process											
3h	Transmit											
NOTES: 1. Command Slot Specific. A 'Yes' in this column indicates the value of the field is independent per Command Slot within a Management Endpoint. A 'No' in this column indicates the same value is reported for either Command Slot.												

Modify a portion of Section 4.2.1.5 (Replay) as follows:

4.2.1.5 Replay

Idle: The Replay Control Primitive requests retransmission of the completion at the offset specified by the RRO field ~~if such a completion is available~~. Following an Abort Control Primitive or a Management Endpoint Reset (refer to section 8.3.3):

- if ~~the a~~ Replay Control Primitive ~~was is~~ received ~~following an Abort Control Primitive or a reset (refer to section 8.3)~~ before any Command Messages are processed, then there is no Response

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Message available to retransmit. The Management Endpoint shall transmit a Response Message with success status with the RR bit cleared to '0'; or

- b) if ~~the a~~ Replay Control Primitive ~~was~~ **is** received following the processing of one or more Command Messages, then the Management Endpoint shall transmit a Response Message with success status with the RR bit set to '1'. The Management Endpoint transmits the MCTP packets associated with the requested Response Message after the Control Primitive Success Response.

Modify a portion of Section 4.2.3 (Management Endpoint Buffer) as follows:

4.2.3 Management Endpoint Buffer

...

~~The contents of the Management Endpoint Buffer are cleared to 0h when the corresponding A~~
Management Endpoint ~~is reset~~ **Reset** (refer to section 8.3.3) shall clear each byte of the corresponding Management Endpoint Buffer to 0h. The contents of the Management Endpoint Buffer are modified by the Management Endpoint Buffer Write command and by Command Messages that generate Response Data with the MEB bit set to '1'.

...

Modify a portion of Figure 67 (SMBus/I2C Frequency – NVMe Management Response) as follows:

5.1.1 SMBus/I2C Frequency (Configuration Identifier 01h)

Figure 67: SMBus/I2C Frequency – NVMe Management Response

Bit	Description												
23:04	Reserved												
03:00	<p>SMBus/I2C Frequency: The current frequency of the SMBus/I2C port. The default value for this field following a reset or power cycle is 1h, if SMBus is supported.</p> <p>A Management Endpoint Reset (refer to section 8.3.3) shall set this field to 1h.</p> <table border="1"> <tr> <th>Value</th><th>Description</th></tr> <tr> <td>0h</td><td>SMBus is not supported or is disabled This value is obsolete for implementations compliant with versions of this specification later than 1.2.</td></tr> <tr> <td>1h</td><td>100 kHz</td></tr> <tr> <td>2h</td><td>400 kHz</td></tr> <tr> <td>3h</td><td>1 MHz</td></tr> <tr> <td>4h to Fh</td><td>Reserved</td></tr> </table>	Value	Description	0h	SMBus is not supported or is disabled This value is obsolete for implementations compliant with versions of this specification later than 1.2.	1h	100 kHz	2h	400 kHz	3h	1 MHz	4h to Fh	Reserved
Value	Description												
0h	SMBus is not supported or is disabled This value is obsolete for implementations compliant with versions of this specification later than 1.2.												
1h	100 kHz												
2h	400 kHz												
3h	1 MHz												
4h to Fh	Reserved												

5.1.3 MCTP Transmission Unit Size (Configuration Identifier 03h)

...

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Figure 69: MCTP Transmission Unit Size – NVMe Management Response

Bits	Description
23:16	Reserved
15:00	MCTP Transmission Unit Size: This field contains the MCTP Transmission Unit Size in bytes to be used by the port. The default value for this field following a reset or power cycle is A Management Endpoint Reset (refer to section 8.3.3) shall cause this field to be set to 40h (64).

5.2.2 Health Status Change (Configuration Identifier 02h)

...

An NVMe Storage Device or NVMe Enclosure supporting the Health Status Change Configuration Identifier in the out-of-band mechanism shall have an independent ~~copy~~ instance of the Composite Controller Status dedicated to the out-of-band mechanism. In the out-of-band mechanism, a Configuration Set command that selects Health Status Change only applies to the ~~copy~~ instance of the Composite Controller Status dedicated to the out-of-band mechanism. Refer to section 5.4 for more details on Composite Controller Status.

An NVMe Storage Device or NVMe Enclosure supporting the Health Status Change Configuration Identifier in the in-band tunneling mechanism shall have an independent ~~copy~~ instance of the Composite Controller Status dedicated to ~~each Controller the in-band tunneling mechanism~~. In the in-band tunneling mechanism, a Configuration Set command that selects Health Status Change only applies to the ~~copy~~ instance of the Composite Controller Status dedicated to the ~~Controller to which the Configuration Set command was issued in-band tunneling mechanism~~.

...

5.3 Controller Health Status Poll

...

An NVMe Storage Device or NVMe Enclosure supporting the Controller Health Status Poll command in the out-of-band mechanism shall have an independent ~~copy~~ instance of the Controller Health Data Structure (refer to Figure 80) and the Controller Health Status Changed Flags (refer to Figure 81) dedicated to the out-of-band mechanism. In the out-of-band mechanism, a Controller Health Status Poll command only applies to the ~~copy~~ instance of the Controller Health Data Structure and the Controller Health Status Changed Flags dedicated to the out-of-band mechanism.

An NVMe Storage Device or NVMe Enclosure supporting the Controller Health Status Poll command in the in-band tunneling mechanism shall have an independent ~~copy~~ instance of the Controller Health Data Structure and the Controller Health Status Changed Flags dedicated to ~~each Controller the in-band tunneling mechanism~~. In the in-band tunneling mechanism, a Controller Health Status Poll command only applies to the ~~copy~~ instance of the Controller Health Data Structure and the Controller Health Status Changed Flags dedicated to the ~~Controller to which the Controller Health Status Poll command was issued in-band tunneling mechanism~~.

...

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Figure 80: Controller Health Data Structure (CHDS)

Description		
Controller Identifier (CTLID): This field specifies the Controller Identifier with which the data contained in this data structure is associated.		
Controller Status (CSTS): This field reports the Controller status.		
Bits	Reset ¹	Description
15:08	0	Reserved
07	HwInit	Firmware Activated (FA): This bit is set to '1' when a new firmware image is activated. Firmware activation is described in the NVM Express Base Specification. The reset value of this bit is set to '1' if a reset caused a new firmware image to be activated.
06	0	Namespace Attribute Changed (NAC): This bit is set to '1' under the same conditions that causes the Namespace Attribute Changed asynchronous event to be sent if Namespace Attribute Notices are enabled as specified in the NVM Express Base Specification. This bit may be set to '1' regardless of whether Namespace Attribute Notices are enabled or not.
05	0 HwInit	Controller Enable Change Occurred (CECO): This bit is set to '1' when the Enable bit (refer to CC.EN in the NVM Express Base Specification) changes state.
04	HwInit	NVM Subsystem Reset Occurred (NSSRO): This bit corresponds to the value of the NVM Subsystem Reset Occurred (refer to CSTS.NSSRO in the NVM Express Base Specification) bit.
03:02	00b	Shutdown Status (SHST): This field corresponds to the value of the Shutdown Status (refer to CSTS.SHST in the NVM Express Base Specification) field.
01	HwInit	Controller Fatal Status (CFS): This bit corresponds to the value of the Controller Fatal Status (refer to CSTS.CFS in the NVM Express Base Specification) bit.
00	0	Ready (RDY): This bit corresponds to the value of the Ready (refer to CSTS.RDY in the NVM Express Base Specification) bit.
...		
NOTES: 1. An NVM Subsystem Reset shall reset the instance of the Controller Status bits dedicated to the out-of-band mechanism and the instance of the Controller Status bits dedicated to each Controller in the NVM Subsystem. The instance of the Controller Status bits dedicated to a Controller shall be reset by a Controller Level Reset (refer to the NVM Express Base Specification) of that Controller. Note that a Controller Level Reset may affect Controller Status bits in the out-of-band mechanism (e.g., a Controller Level Reset causes the CECO bit in the instance of the Controller Status bits dedicated to the out-of-band mechanism to be set to '1'). No instance of the Controller Status bits shall be reset by any other resets other than the resets documented by this note.		

...

5.3.2 Filtering by Controller Health Status Changed Flags


...

Figure 81: Controller Health Status Changed Flags (CHSCF)

Bits	Reset ¹	Description
15:13	0	Reserved
12	0	Critical Warning (CWARN): This bit is set to '1' when any of the Critical Warning bits in the Controller Health Data Structure transition from '0' to '1'.

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Figure 81: Controller Health Status Changed Flags (CHSCF)

Bits	Reset ¹	Description
11	0	Available Spare (SPARE): This bit is set to '1' when the Available Spare field in the Controller Health Data Structure changes state.
10	0	Percentage Used (PDLU): This bit is set to '1' when the Percentage Used field in the Controller Health Data Structure changes state.
09	0	Composite Temperature Change (CTEMP): This bit is set to '1' when the Composite Temperature field in the Controller Health Data Structure changes state.
08	HwInit	Controller Status Change (CSTS): This bit is set to '1' when the Shutdown Status field in the Controller Health Data Structure changes state or when the Ready, Controller Fatal Status, NVM Subsystem Reset Occurred, Controller Enable Change Occurred, Namespace Attribute Changed, or Firmware Activated bit in the Controller Health Data Structure transitions from '0' to '1'.
07	HwInit	Firmware Activated (FA): This bit is set to '1' when the Firmware Activated bit in the Controller Health Data Structure transitions from '0' to '1'.
06	0	Namespace Attribute Changed (NAC): This bit is set to '1' when the Namespace Attribute Changed bit in the Controller Health Data Structure transitions from '0' to '1'.
05	 HwInit	Controller Enable Change Occurred (CECO): This bit is set to '1' when the Controller Enable Change Occurred bit in the Controller Health Data Structure transitions from '0' to '1'.
04	HwInit	NVM Subsystem Reset Occurred (NSSRO): This bit is set to '1' when the NVM Subsystem Reset Occurred bit in the Controller Health Data Structure transitions from '0' to '1'.
03	0	Reserved
02	0	Shutdown Status (SHST): This bit is set to '1' when the Shutdown Status field in the Controller Health Data Structure changes state.
01	HwInit	Controller Fatal Status (CFS): This bit is set to '1' when the Controller Fatal Status bit in the Controller Health Data Structure transitions from '0' to '1'.
00	0	Ready (RDY): This bit is set to '1' when the Ready bit in the Controller Health Data Structure transitions from '0' to '1'.
<p>NOTES:</p> <p>1. An NVM Subsystem Reset shall reset the instance of the Controller Status bits dedicated to the out-of-band mechanism and the instance of the Controller Status bits dedicated to each Controller in the NVM Subsystem.</p> <p>The instance of the Controller Status bits dedicated to a Controller shall be reset by a Controller Level Reset (refer to the NVM Express Base Specification) of that Controller. Note that a Controller Level Reset may affect Controller Status bits in the out-of-band mechanism (e.g., a Controller Level Reset causes the CECO bit in the instance of the Controller Status bits dedicated to the out-of-band mechanism to be set to '1').</p> <p>No instance of the Controller Status bits shall be reset by any other resets other than the resets documented by this note.</p>		

Modify a portion of Section 5.6 (NVM Subsystem Health Status Poll) as follows:

5.6 NVM Subsystem Health Status Poll

...

An NVMe Storage Device or NVMe Enclosure supporting the NVM Subsystem Health Status Poll command using the out-of-band mechanism shall have an independent ~~copy~~ instance of the NVM Subsystem Health Data Structure (refer to Figure 90) dedicated to the out-of-band mechanism. In the out-of-band mechanism, an NVM Subsystem Health Status Poll command only applies to the ~~copy~~ instance of the NVM Subsystem Health Data Structure dedicated to the out-of-band mechanism.

An NVMe Storage Device or NVMe Enclosure supporting the NVM Subsystem Health Status Poll command using the in-band tunneling mechanism shall have an independent ~~copy~~ instance of the NVM Subsystem Health Data Structure dedicated to ~~each Controller the in-band tunneling mechanism~~. In the in-band tunneling mechanism, an NVM Subsystem Health Status Poll command only applies to the ~~copy~~ instance of the NVM Subsystem Health Data Structure dedicated to the ~~Controller to which the NVM Subsystem Health Poll command was issued in-band tunneling mechanism~~.

...

Figure 90: NVM Subsystem Health Data Structure (NSHDS)

Description	
NVM Subsystem Status (NSS): This field indicates the status of the NVM Subsystem.	
Bit	Description
7:6	Reserved
5	Drive Functional (DF): This bit is set to '1' to indicate an NVM Subsystem is functional. If cleared to '0', then there is an unrecoverable failure detected in the NVM Subsystem.
4	Reset Not Required (RNR): This bit is set to '1' to indicate the NVM Subsystem does not need a reset to resume normal operation. If this bit is cleared to '0', then the NVM Subsystem has experienced an error that prevents continued normal operation. An NVM Subsystem ResetA Controller Level Reset is required to resume normal operation. If this bit is set to '1', then the NVM Subsystem has not experienced an error that prevents continued normal operation. An NVM Subsystem Reset is not required to resume normal operation.
3	Port 0 PCIe Link Active (P0LA): This bit is set to '1' to indicate the first port's PCIe link is up (i.e., the Data Link Control and Management State Machine is in the DL_Active state). If cleared to '0', then the PCIe link is down.
2	Port 1 PCIe Link Active (P1LA): This bit is set to '1' to indicate the second port's PCIe link is up. If cleared to '0', then the second port's PCIe link is down or not present.
1:0	Reserved
...	

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.


Figure 90: NVM Subsystem Health Data Structure (NSHDS)

Description

5:4

Composite Controller Status (CCS): This field reports the composite status of all Controllers in the NVM Subsystem.

The bits in this field are cleared after the NVM Subsystem Health Data Structure (refer to Figure 90) is returned in a Success Response associated with an NVM Subsystem Health Status Poll command where the Clear Status bit set. A Configuration Set command that selects Health Status Change may be used to clear selected bits to '0'.

Bits	Reset ¹	Description
15:13	0	Reserved
12	0	Critical Warning (CWARN): This bit is set to '1' when the Critical Warning bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
11	0	Available Spare (SPARE): This bit is set to '1' when the Available Spare bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
10	0	Percentage Used (PDLU): This bit is set to '1' when the Percentage Used field in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
09	0	Composite Temperature Change (CTEMP): This bit is set to '1' when the Composite Temperature field in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
08	Hwlnit	Controller Status Change (CSTS): This bit is set to '1' when the Controller Status field in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
07	Hwlnit	Firmware Activated (FA): This bit is set to '1' when the Firmware Activated bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
06	0	Namespace Attribute Changed (NAC): This bit is set to '1' when the Namespace Attribute Changed bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
05	 Hwlnit	Controller Enable Change Occurred (CECO): This bit is set to '1' when the Controller Enable Change Occurred bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
04	Hwlnit	NVM Subsystem Reset Occurred (NSSRO): This bit is set to '1' when the value of the NVM Subsystem Reset Occurred (CSTS.NSSRO) bit transitions from a '0' to a '1' in one or more Controllers in the NVM Subsystem.
03	0	Reserved
02	0	Shutdown Status (SHST): This bit is set to '1' when the Shutdown Status bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
01	Hwlnit	Controller Fatal Status (CFS): This bit is set to '1' when the Controller Fatal Status bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.
00	0	Ready (RDY): This bit is set to '1' when the Ready bit in the Controller Health Status Changed Flags transitions from '0' to '1' in one or more Controllers in the NVM Subsystem.

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Figure 90: NVM Subsystem Health Data Structure (NSHDS)

Description
<p>NOTES:</p> <p>1. An NVM Subsystem Reset shall reset the instance of the Controller Status bits dedicated to the out-of-band mechanism and the instance of the Controller Status bits dedicated to each Controller in the NVM Subsystem.</p> <p>The instance of the Controller Status bits dedicated to a Controller shall be reset by a Controller Level Reset (refer to the NVM Express Base Specification) of that Controller. Note that a Controller Level Reset may affect Controller Status bits in the out-of-band mechanism (e.g., a Controller Level Reset causes the CECO bit in the instance of the Controller Status bits dedicated to the out-of-band mechanism to be set to '1').</p> <p>No instance of the Controller Status bits shall be reset by any other resets other than the resets documented by this note.</p>

Modify a portion of Section 5.8 (Reset) as follows:

5.8 Reset

The Reset command ~~may be~~ is used to initiate ~~a~~ the reset defined by the Reset Type field.

...

Modify a portion of Section 8.1 (Out-of-Band Operational Times) as follows:

8.1 Out-of-Band Operational Times

Figure 146: Operations Supported During NVM Subsystem Power States

Operation ^{3, 4}	Powered Off -All Power Rails Off	Powered On -All Power Rails On	Auxiliary Power Only ² -Main Power Off -Auxiliary Power On	Main Power Only ² -Main Power On -Auxiliary Power Off
SMBus/I2C VPD and SMBus/I2C Mux Access	Not Supported	Supported	Supported	Implementation Specific
SMBus/I2C MCTP Access	Not Supported	Supported	Optional ¹	Implementation Specific
PCIe MCTP Access	Not Supported	Supported	Not Supported	Supported

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

Figure 146: Operations Supported During NVM Subsystem Power States

Operation ^{3, 4}	Powered Off -All Power Rails Off	Powered On -All Power Rails On	Auxiliary Power Only ² -Main Power Off -Auxiliary Power On	Main Power Only ² -Main Power On -Auxiliary Power Off
NOTES: 1. An implementation that supports SMBus/I2C MCTP Access during Auxiliary Power may support a subset of commands during this power state. The commands that are supported are implementation specific. 2. Auxiliary Power Only and Main Power Only columns are not applicable to form factors that do not define Auxiliary power. 3. An SMBus Reset may prevent access as described in section 8.3.4. A PCIe Reset may prevent access as described in 8.3.TBD. 4. Firmware activation may impact access as described in section 8.3.2.				

Modify a portion of Section 8.2 (Vital Product Data) as follows:

8.2 Vital Product Data

...

Figure 148 shows an I2C Read where the A6h addresses and Command Offset are provided by the Management Controller followed by data being returned from the Management Endpoint. The Command Offset as shown in Figure 148 is stored internal to the NVMe Storage Device (i.e., the internal offset).

If an I2C Read is issued, then data is returned from the internal offset within the FRU Information Device and then the internal offset is incremented by 1h. If the Management Controller reads the last byte of the FRU Information Device (refer to Maximum FRU Information Size) via an I2C Read, then the internal offset shall be cleared to 0h (i.e., rolls over to 0h). If only one byte of the Command Offset is provided by the Management Controller, then the least-significant byte of the internal offset shall be set to that value and the most-significant byte of the internal offset shall be cleared to 0h.

The internal offset shall be cleared to 0h ~~following by~~ a power ~~on cycle~~ of the FRU Information Device. ~~Implementations are allowed to maintain the current internal offset value or~~ All other SMBus Resets should clear the internal offset value ~~it~~ to 0h ~~following a reset of the FRU Information Device.~~

Modify Section 8.3 (Reset) as follows:

8.3 Reset Architecture

This section describes the reset architecture defined by this specification that ~~are~~ is applicable to NVMe Storage Devices and NVMe Enclosures. Additional requirements and recommendations for resets are specified elsewhere in this specification.

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

8.3.1 NVM Subsystem Reset

An NVM Subsystem Reset is initiated under the conditions outlined in the NVM Express Base Specification (e.g., when main power is applied to the NVM Subsystem) and section 5.8 of this specification. ~~In addition to these conditions, if NVM Subsystem Reset is supported, then it may be initiated by processing a Reset command.~~

An NVM Subsystem Reset initiated via the out-of-band mechanism may interfere with host software. A Management Controller should coordinate with the host. Coordination between a Management Controller and a host are outside the scope of this specification.

When an NVM Subsystem Reset is initiated, the entire NVM Subsystem ~~is~~ shall be reset. This includes all NVM Subsystem ports (PCIe and SMBus/I2C), SMBus/I2C elements (e.g., SMBus/I2C Management Endpoints, FRU Information Devices, SMBus/I2C Muxes, etc.), PCIe VDM Management Endpoints, and Controller Management Interfaces. If an NVMe Storage Device does not contain an SMBus/I2C port, then a NVM Subsystem Reset should reset the FRU Information Device if the FRU Information Device supports a reset mechanism. ~~All~~ On an NVM Subsystem Reset, any internal state ~~is~~ of the NVM Subsystem should be returned to its power-on ~~default~~ condition.

8.3.2 Controller Level Reset

A Controller Level Reset is initiated under the conditions outlined in the NVM Express Base Specification.

A Controller Level Reset initiated via the out-of-band mechanism may interfere with host software. A Management Controller should coordinate with the host. Coordination between a Management Controller and a host are outside the scope of this specification.

The actions performed on a Controller Level Reset are outlined in the NVM Express Base Specification. A Controller Level Reset has no effect on the Controller Management Interface associated with that Controller, the PCI Express port associated with that Controller, or a Management Endpoint associated with that port. A Controller Level Reset ~~also does~~ shall not stop the servicing of the Management Interface Command Set, ~~or~~ NVM Express Admin Command Set commands, or Control Primitives that target that Controller (i.e., the NVM Express Admin Command Set is still available even though the NVMe Controller may be disabled or held in reset) ~~or Control Primitives~~.

A Controller Level Reset may ~~affect~~ prevent PCIe Commands ~~Set commands~~ from being processed on that Controller (refer to section 8.1). If a PCIe Command is ~~affected~~ prevented from being processed due to a Controller Level Reset, then ~~the~~ that ~~e~~PCIe Command shall be ~~is~~ completed with status PCIe Inaccessible.

On a Controller Level Reset, any internal state of the Controller should be returned to its power-on condition.

A Controller Level Reset that causes a new firmware image to activate is considered a special event and may impact the operation of the Controller Management Interface associated with one or more Controllers, servicing of NVMe-MI Messages, or Management Endpoints within an NVM Subsystem. This impact is unspecified and vendor specific. The Management Controller and host should coordinate the activation of a new firmware image. Coordination between a Management Controller and a host are outside the scope of this specification.

Additional requirements and recommendations for Controller Level Resets are specified elsewhere in this specification. For example, bits and fields that are dedicated to each Controller in the in-band tunneling mechanism are reset as defined in Figure 80, Figure 81, and Figure 90.

Technical input submitted to NVM Express® is subject to the terms of the NVM Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

8.3.3 Management Endpoint Reset

~~A~~The following shall cause a Management Endpoint Reset ~~is initiated~~:

- an NVM Subsystem Reset of the NVM Subsystem containing the Management Endpoint; or
- ~~under~~ the conditions for resetting an MCTP endpoint outlined in the MCTP Base Specification or the associated MCTP transport binding specifications.

In addition to these conditions, a Management Endpoint Reset shall be initiated on the Management Endpoint associated with a PCI Express port ~~is reset~~ when ~~that~~ that PCI Express port ~~is in a PCI Express conventional reset state~~ undergoes a PCIe Reset (refer to section 8.3.TBD) or is powered on. A Management Endpoint Reset shall be initiated on the Management Endpoint associated with an SMBus/I2C port when that SMBus/I2C port undergoes an SMBus Reset (refer to section 8.3.4) or is powered on.

~~When~~If a Management Endpoint Reset is initiated, then:

- each Command Slot in that Management Endpoint shall behave as if an implicit Abort Control Primitive (refer to section 4.2.1.3) was received with the exception that the Management Endpoint shall not transmit any Abort Control Primitive Response Messages;
- any Control Primitives being processed by that Management Endpoint shall be dropped (silently discarded); and
- ~~the any internal~~ state of that Management Endpoint ~~should be~~ is returned to its default power-on condition ~~and any commands associated with that Management Endpoint being processed are aborted~~.

~~A reset of a~~ Management Endpoint Reset of a Management Endpoint in an NVM Subsystem shall not affect any other Management Endpoint in the NVM Subsystem or any other NVM Subsystem entity in the NVM Subsystem. Note that for implementations compliant to version 1.1 and earlier of this specification, implementations may block MCTP accesses on additional Management Endpoints during a PCI Express conventional reset of a PCIe VDM Management Endpoint. ~~MCTP accesses may not be supported on other PCIe or SMBus/I2C Management Endpoints in the NVM Subsystem.~~

Additional requirements and recommendations for Management Endpoint Resets are specified elsewhere in this specification. For example, a Management Endpoint Reset:

- resets bits and fields that are dedicated to the out-of-band mechanism as defined in Figure 80, Figure 81, and Figure 90;
- resets the value of the MCTP Transmission Unit Size field as defined by Figure 69; and
- clears the Control Primitive Specific Response field to 0h as defined in Figure 42.

8.3.4 SMBus Reset

SMBus clock-low recovery is the ability to reset communication when the SMBus/I2C clock is low for longer than $t_{\text{TIMEOUT,MIN}}$ (refer to the SMBus Specification). SMBus/I2C Management Endpoints shall support SMBus clock-low recovery. It is strongly recommended that any SMBus/I2C element other than the SMBus/I2C Management Endpoint (refer to Figure 16 for a list of SMBus/I2C elements) should support SMBus clock-low recovery. SMBus clock-low recovery shall cause an SMBus Reset. An SMBus Reset caused by SMBus clock-low recovery shall not cause ARP-assigned addresses to be reset to their default values. ~~All SMBus/I2C elements should support the recommendation for SMBus Reset when the SMBus/I2C clock is low for longer than $t_{\text{TIMEOUT,MIN}}$ (refer to the SMBus Specification).~~

Some form factors may also specify one or more form factor-specific ~~separate SMBus Reset~~ mechanisms to reset the SMBus (e.g., SMBRST# as defined in SFF-TA-1009 or the rising edge of the +3.3 Vaux rail as defined in the PCI Express SFF-8639 Module Specification). ~~If such mechanisms are~~ Any form factor-specific mechanisms to reset the SMBus supported by an NVMe Storage Device or NVMe Enclosure NVM Subsystem, then the NVM Subsystem shall cause an SMBus Reset. ~~propagate the reset to all~~

Technical input submitted to NVMe Express® is subject to the terms of the NVMe Express® Participant's agreement. Copyright © 2008 to 2022 NVMe® Corporation.

~~SMBus/I2C elements on the NVM Subsystem and translate the reset, if needed, to Expansion Connector form factors.~~

An SMBus Reset shall cause a Management Endpoint Reset of the SMBus/I2C Management Endpoint. For any SMBus/I2C element other than the SMBus/I2C Management Endpoint (refer to [Figure 16](#) for a list of SMBus/I2C elements), it is strongly recommended that an SMBus Reset should reset that SMBus/I2C element.

An SMBus Reset shall cause an SMBus reset mechanism defined for the Expansion Connector to be applied to each Expansion Connector in the NVMe Storage Device.

~~If the an SMBus/I2C port element on an NVM Subsystem is transmitting a Response Message, then an SMBus Reset shall cause the SMBus/I2C port it to attempt to generate a STOP condition as defined in (refer to the SMBus Specification) within 5 ms from the assertion of SMBus Reset or after the current data byte in the transfer process. The NVM Subsystem shall remain idle on SMBus/I2C port shall remain in the bus idle condition (refer to the SMBus Specification) for the remainder of the SMBus Reset assertion even if other SMBus/I2C elements a Management Controller attempts to address it access the SMBus/I2C port. An NVM Subsystem SMBus/I2C port shall support SMBus/I2C accesses starting from the de-assertion of SMBus Reset within the same timing constraints as are applicable to transitioning from an unsupported to a supported power state as defined in section 8.1 be ready to receive a START condition as defined in the SMBus Specification within 10 ms after SMBus Reset de-assertion.~~

~~An SMBus Reset shall not modify ARP assigned addresses. Management Controllers may send an ARP reset after the SMBus Reset if addresses are to be reinitialized.~~

~~An SMBus Reset shall be treated by each Command Slot in the SMBus/I2C Management Endpoint as if an implicit Abort Control Primitive (refer to section 4.2.1.3) was received with the exception that the Management Endpoint does not transmit the Abort Control Primitive Response Messages.~~

Additional requirements and recommendations for SMBus Resets are specified elsewhere in this specification. For example, and SMBus Reset:

- resets the value of the SMBus/I2C Frequency field as defined in [Figure 67](#); and
- should clear the internal Command Offset for the VPD to 0h as defined in section [8.2](#).

8.3. [TBD](#) PCIe Reset

A PCIe Reset is generated by a Conventional Reset (refer to the PCI Express Base Specification) or a Function Level Reset (refer to the PCI Express Base Specification). PCIe Resets have additional impacts on in-band traffic and NVMe Controller operations which are outside the scope of this specification.

A PCIe Reset shall cause a Management Endpoint Reset of the PCIe VDM Management Endpoint associated with the PCI Express port being reset.

A PCIe VDM Management Endpoint shall support PCIe MCTP accesses after a PCIe Reset is de-asserted within the same timing constraints as are applicable to transitioning from an unsupported to a supported NVM Subsystem power state as defined by section [8.1](#).